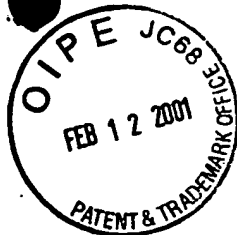


862.C2001



PATENT APPLICATION

2641
#2
TR
2/22/01

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of:

YASHUHIRO KOMORI, ET AL.

Application No.: 09/661,394

Filed: September 13, 2000

For: SPEECH INPUT
TERMINAL, SPEECH
RECOGNITION
APPARATUS, SPEECH
COMMUNICATION SYSTEM,
AND SPEECH
COMMUNICATION METHOD

Examiner: UNASSIGNED

Group Art Unit: UNASSIGNED

February 12, 2001

RECEIVED
FEB 16 2001
TC 2600 MAILROOM

Commissioner for Patents
Washington, D.C. 20231

CLAIM TO PRIORITY

Sir:

Applicants hereby claim priority under the International Convention and all rights
to which they are entitled under 35 U.S.C. § 119 based upon the following Japanese

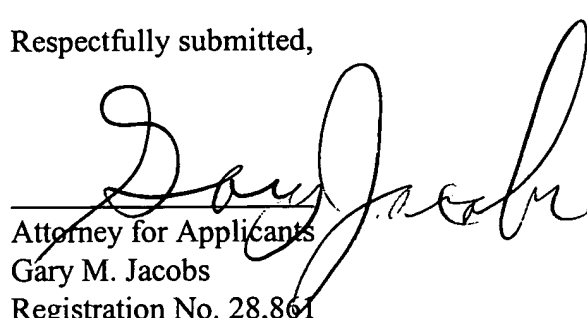
Priority Application No.:

11-260760, filed September 14, 1999.

A certified copy of the priority document is enclosed.

Applicants' undersigned attorney may be reached in our Washington, D.C. office by telephone at (202) 530-1010. All correspondence should continue to be directed to our below-listed address.

Respectfully submitted,

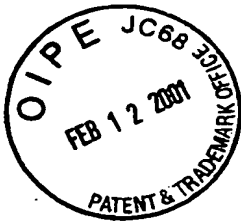


Attorney for Applicants
Gary M. Jacobs
Registration No. 28,861

FITZPATRICK, CELLA, HARPER & SCINTO
30 Rockefeller Plaza
New York, New York 10112-3801
Facsimile No.: (212) 218-2200

09/861,344
Yasuhito Komori
9/13/00

(translation of the front page of the priority document of
Japanese Patent Application No. 11-260760)



PATENT OFFICE
JAPANESE GOVERNMENT

This is to certify that the annexed is a true copy of the
following application as filed with this Office.

RECEIVED
FEB 16 2001
TC 2600 MAIL ROOM

Date of Application: September 14, 1999

Application Number : Patent Application 11-260760

Applicant(s) : Canon Kabushiki Kaisha

October 6, 2000

Commissioner,
Patent Office

Kouzo OIKAWA

Certification Number 2000-3082056

CFM 2001 25

日 本 国 特 許 庁
PATENT OFFICE
JAPANESE GOVERNMENT

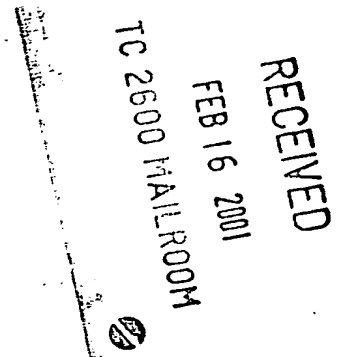
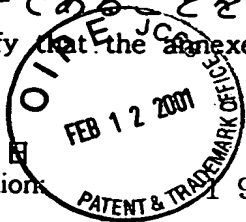
別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日
Date of Application: 1999年 9月14日

出 願 番 号
Application Number: 平成11年特許願第260760号

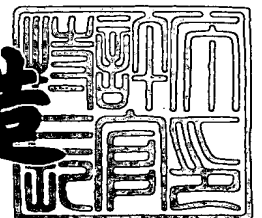
出 願 人
Applicant (s): キヤノン株式会社



2000年10月 6日

特許庁長官
Commissioner,
Patent Office

及川耕造



出証番号 出証特2000-3082056

【書類名】 特許願

【整理番号】 3808025

【提出日】 平成11年 9月14日

【あて先】 特許庁長官殿

【国際特許分類】 G06F 3/00

【発明の名称】 音声入力端末器、音声認識装置、音声通信システム及び
音声通信方法

【請求項の数】 32

【発明者】

 【住所又は居所】 東京都大田区下丸子3丁目30番2号 キヤノン株式会
社内

 【氏名】 小森 康弘

【発明者】

 【住所又は居所】 東京都大田区下丸子3丁目30番2号 キヤノン株式会
社内

 【氏名】 山田 雅章

【特許出願人】

 【識別番号】 000001007

 【氏名又は名称】 キヤノン株式会社

【代理人】

 【識別番号】 100076428

 【弁理士】

 【氏名又は名称】 大塚 康徳

 【電話番号】 03-5276-3241

【選任した代理人】

 【識別番号】 100093908

 【弁理士】

 【氏名又は名称】 松本 研一

 【電話番号】 03-5276-3241

【選任した代理人】

【識別番号】 100101306

【弁理士】

【氏名又は名称】 丸山 幸雄

【電話番号】 03-5276-3241

【手数料の表示】

【予納台帳番号】 003458

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9704672

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 音声入力端末器、音声認識装置、音声通信システム及び音声通信方法

【特許請求の範囲】

【請求項 1】 有線又は無線の通信網を介して音声認識装置へ音声データを伝送する音声入力端末器であって、

音声の入力手段と、

音声認識のための前記音声入力端末器固有の又は利用状況の情報を作成する手段と、

当該情報を前記音声認識装置へ送信する通信手段と、
を備えたことを特徴とする音声入力端末器。

【請求項 2】 前記情報が、前記音声入力手段の特性、雑音特性、又は、話者特性、の少なくともいずれか一つに基づく情報であることを特徴とする請求項 1 に記載の音声入力端末器。

【請求項 3】 前記音声認識装置から、前記情報に基づく通信のためのデータ変換条件を受け取った場合に、当該変換条件に基づいて前記音声データを変換する手段を備えたことを特徴とする請求項 1 に記載の音声入力端末器。

【請求項 4】 前記情報を保存する手段と、
各通信時において前記情報に変更があったか否かを判定する手段と、
前記情報に変更がなかった場合は、その旨を前記音声認識装置へ通知する手段と、
を備えたことを特徴とする請求項 1 に記載の音声入力端末器。

【請求項 5】 前記情報に基づいて、音声認識モデルを作成する手段を備え、
前記通信手段は、前記情報及び／又は当該音声認識モデルを、前記音声認識装置へ送信することを特徴とする請求項 1 に記載の音声入力端末器。

【請求項 6】 有線又は無線の通信網を介して音声入力端末器から伝送される音声データに対して音声認識処理を実行する音声認識装置であって、
前記音声入力端末器から、音声認識のための前記音声入力端末器固有の又は利

用状況の情報を受け取った場合に、当該情報に基づいて音声認識処理を実行する手段を備えたことを特徴とする音声認識装置。

【請求項 7】 有線又は無線の通信網を介して音声入力端末器から伝送される音声データに対して音声認識処理を実行する音声認識装置であって、

伝送された音声データに基づいて、音声認識のための前記音声入力端末器固有の又は利用状況の情報を作成する手段と、

当該情報に基づいて音声認識処理を実行する手段と、
を備えたことを特徴とする音声認識装置。

【請求項 8】 前記情報に基づいて音声認識モデルを作成する手段を備えたことを特徴とする請求項 6 又は 7 に記載の音声認識装置。

【請求項 9】 有線又は無線の通信網を介して音声入力端末器から伝送される音声データに対して音声認識処理を実行する音声認識装置であって、

前記音声入力端末器から、音声認識のための前記音声入力端末器固有の又は利用状況の情報を受け取った場合に、前記情報に基づいて、通信のためのデータ変換条件を決定する手段と、

当該データ変換条件を前記音声入力端末器へ送出する手段と、を備えたことを特徴とする音声認識装置。

【請求項 10】 有線又は無線の通信網を介して音声入力端末器から伝送される音声データに対して音声認識処理を実行する音声認識装置であって、

伝送された音声データに基づいて、音声認識のための前記音声入力端末器固有の又は利用状況の情報を作成する手段と、

前記情報に基づいて、通信のためのデータ変換条件を決定する手段と、

当該データ変換条件を前記音声入力端末器へ送出する手段と、
を備えたことを特徴とする音声認識装置。

【請求項 11】 前記データ変換条件が、前記情報に基づいて作成された量子化テーブルに基づくデータ変換条件を含むことを特徴とする請求項 9 又は 10 に記載の音声認識装置。

【請求項 12】 前記音声入力端末器が複数の場合、各々の前記音声入力端末器毎に、前記情報を保存する手段を備えたことを特徴とする請求項 6、7、9

又は 10 のいずれか 1 項に記載の音声認識装置。

【請求項 13】 前記音声入力端末器が複数の場合、各々の前記音声入力端末器毎に、前記音声認識モデルを保存する手段を備えたことを特徴とする請求項 8 に記載の音声認識装置。

【請求項 14】 前記音声入力端末器が複数の場合、各々の前記音声入力端末器毎に、前記データ変換条件を保存する手段を備えたことを特徴とする請求項 9 又は 10 に記載の音声認識装置。

【請求項 15】 有線又は無線の通信網を介して通信可能な音声入力端末器と音声認識装置とからなる音声通信システムであって、

前記音声入力端末器は、

音声入力手段と、

音声認識のための前記音声入力端末器固有の又は利用状況の情報を作成する手段と、

当該情報を前記音声認識装置へ送信する通信手段と、を備え、

前記音声認識装置は、

前記情報に基づいて音声認識処理を実行する手段を備えたことを特徴とする音声通信システム。

【請求項 16】 有線又は無線の通信網を介して通信可能な音声入力端末器と音声認識装置とからなる音声通信システムであって、

前記音声認識装置は、

前記音声入力端末器からの音声データに基づいて、音声認識のための前記音声入力端末器固有の又は利用状況の情報を作成する手段と、

前記情報に基づいて音声認識処理を実行する手段と、
を備えたことを特徴とする音声通信システム。

【請求項 17】 有線又は無線の通信網を介して通信可能な音声入力端末器と音声認識装置とからなる音声通信システムであって、

前記音声入力端末器は、

音声入力手段と、

音声認識のための前記音声入力端末器固有の又は利用状況の情報を作成する

手段と、

当該情報を前記音声認識装置へ送信する通信手段と、を備え、

前記音声認識装置は、

前記情報に基づいて、通信のためのデータ変換条件を決定する手段と、

当該データ変換条件を前記音声入力端末器に送信する手段と、

を備えたことを特徴とする音声通信システム。

【請求項 1 8】 有線又は無線の通信網を介して通信可能な音声入力端末器と音声認識装置とからなる音声通信システムであって、

前記音声認識装置は、

前記音声入力端末器からの音声データに基づいて、音声認識のための前記音声入力端末器固有の又は利用状況の情報を作成する手段と、

前記情報に基づいて、通信のためのデータ変換条件を決定する手段と、

当該データ変換条件を前記音声入力端末器に送信する手段と、

を備えたことを特徴とする音声通信システム。

【請求項 1 9】 有線又は無線の通信網を介して、音声入力端末器から音声認識装置へ音声データを伝送する音声通信方法であって、

前記音声入力端末器において、

音声認識のための前記音声入力端末器固有の又は利用状況の情報を作成する工程と、

当該情報を前記音声認識装置へ送信する通信工程と、

を含むことを特徴とする音声通信方法。

【請求項 2 0】 有線又は無線の通信網を介して音声入力端末器から伝送される音声データに対して音声認識処理を実行する音声通信方法であって、

前記音声入力端末器から、音声認識のための前記音声入力端末器固有の又は利用状況の情報を受け取った場合に、当該情報に基づいて音声認識処理を実行する工程を含むことを特徴とする音声通信方法。

【請求項 2 1】 有線又は無線の通信網を介して音声入力端末器から伝送される音声データに対して音声認識処理を実行する音声通信方法であって、

前記音声入力端末器から伝送されたデータに基づいて、音声認識のための前記

音声入力端末器固有の又は利用状況の情報を作成する工程と、

当該情報に基づいて音声認識処理を実行する工程と、
を含むことを特徴とする音声通信方法。

【請求項 2 2】 有線又は無線の通信網を介して音声入力端末器から伝送される音声データに対して音声認識処理を実行する音声通信方法であって、

前記音声入力端末器から、音声認識のための前記音声入力端末器固有の又は利用状況の情報を受け取った場合に、当該情報に基づいて通信のためのデータ変換条件を決定する工程と、

当該データ変換条件を前記音声入力端末器へ送出する工程と、
を含むことを特徴とする音声通信方法。

【請求項 2 3】 有線又は無線の通信網を介して音声入力端末器から伝送される音声データに対して音声認識処理を実行する音声通信方法であって、

前記音声入力端末器から伝送されたデータに基づいて、音声認識のための前記音声入力端末器固有の又は利用状況の情報を作成する工程と、

当該情報に基づいて通信のためのデータ変換条件を決定する工程と、
当該データ変換条件を前記音声入力端末器へ送出する工程と、
を含むことを特徴とする音声通信方法。

【請求項 2 4】 有線又は無線の通信網を介して通信可能な音声入力端末器と音声認識装置との間での音声通信方法であって、

前記音声入力端末器において、
音声認識のための前記音声入力端末器固有の又は利用状況の情報を作成する工程と、

当該情報を前記音声認識装置へ送信する通信工程と、を含み、
前記音声認識装置において、

前記情報に基づいて音声認識処理を実行する工程、を含むことを特徴とする音声通信方法。

【請求項 2 5】 有線又は無線の通信網を介して通信可能な音声入力端末器と音声認識装置との間での音声通信方法であって、

前記音声認識装置において、

前記音声入力端末器からの音声データに基づいて、音声認識のための前記音声入力端末器固有の又は利用状況の情報を作成する工程と、

前記情報に基づいて音声認識処理を実行する工程と、
を含むことを特徴とする音声通信方法。

【請求項 2 6】 有線又は無線の通信網を介して通信可能な音声入力端末器と音声認識装置との間での音声通信方法であって、

前記音声入力端末器において、
音声認識のための前記音声入力端末器固有の又は利用状況の情報を作成する工程と、

当該情報を前記音声認識装置へ送信する通信工程と、を含み、
前記音声認識装置において、

前記情報に基づいて、通信のためのデータ変換条件を決定する工程と、
当該データ変換条件を前記音声入力端末器に送信する工程と、
を含むことを特徴とする音声通信方法。

【請求項 2 7】 有線又は無線の通信網を介して通信可能な音声入力端末器と音声認識装置との間での音声通信方法であって、

前記音声認識装置において、
前記音声入力端末器からの音声データに基づいて、音声認識のための前記音声入力端末器固有の又は利用状況の情報を作成する工程と、

前記情報に基づいて、通信のためのデータ変換条件を決定する工程と、
当該データ変換条件を前記音声入力端末器に送信する工程と、
を含むことを特徴とする音声通信方法。

【請求項 2 8】 有線又は無線の通信網を介して音声入力端末器から音声認識装置へ音声データを伝送するために、コンピュータを、

音声認識のための前記音声入力端末器固有の又は利用状況の情報を作成する手段、

当該情報を前記音声認識装置へ送信する通信手段、
として機能させるプログラムを記録した記憶媒体。

【請求項 2 9】 有線又は無線の通信網を介して音声入力端末器からの音声

データについて音声認識処理を実行するために、コンピュータを、

前記音声入力端末器から、音声認識のための前記音声入力端末器固有の又は利用状況の情報を受け取った場合に、当該情報に基づいて音声認識処理を実行する手段、

として機能させるプログラムを記録した記憶媒体。

【請求項 3 0】 有線又は無線の通信網を介して音声入力端末器からの音声データについて音声認識処理を実行するために、コンピュータを、

前記音声入力端末器から伝送されたデータに基づいて、音声認識のための前記音声入力端末器固有の又は利用状況の情報を作成する手段、

当該情報に基づいて音声認識処理を実行する手段、
として機能させるプログラムを記録した記憶媒体。

【請求項 3 1】 有線又は無線の通信網を介して音声入力端末器からの音声データについて音声認識処理を実行するために、コンピュータを、

前記音声入力端末器から、音声認識のための前記音声入力端末器固有の又は利用状況の情報を受け取った場合に、当該情報に基づいて通信のためのデータ変換条件を決定する手段、

当該データ変換条件を前記音声入力端末器へ送出する手段、
として機能させるプログラムを記録した記憶媒体。

【請求項 3 2】 有線又は無線の通信網を介して音声入力端末器からの音声データについて音声認識処理を実行するために、コンピュータを、

前記音声入力端末器から伝送されたデータに基づいて、音声認識のための前記音声入力端末器固有の又は利用状況の情報を作成する手段、

当該情報に基づいて通信のためのデータ変換条件を決定する手段、

当該データ変換条件を前記音声入力端末器へ送出する手段、
として機能させるプログラムを記録した記憶媒体。

【発明の詳細な説明】

【0 0 0 1】

【発明の属する技術分野】

本発明は、通信網を介して、音声データを伝送し、その音声認識を実行するた

めの、音声入力端末器、音声認識装置、音声通信システム及び音声通信方法に関する。

【0002】

【従来の技術】

携帯電話等の音声入力端末器から、通信網を介して音声データをホストサーバへ送出し、特定の情報を引き出す処理等を実行する音声通信システムが提案されている。係る音声通信システムでは、音声によりデータの送受信が可能であるため、操作が簡単であるという利点がある。

【0003】

【発明が解決しようとする課題】

しかし、携帯電話等の音声入力端末器自体の特性や、その周辺環境等により、音声データが変動して十分な音声認識ができない場合がある。

【0004】

また、いかなる場合も同じ通信条件により通信を行うため、必ずしも通信効率がよいとはいえなかった。

【0005】

従って、本発明の目的は、音声入力端末器に関する事情に基づいて、最適な音声認識又は通信を達成し得る音声入力端末器、音声認識装置、音声通信システム及び音声通信方法を提供することにある。

【0006】

【課題を解決するための手段】

本発明によれば、有線又は無線の通信網を介して音声認識装置へ音声データを伝送する音声入力端末器であって、音声の入力手段と、音声認識のための前記音声入力端末器固有の又は利用状況の情報を作成する手段と、当該情報を前記音声認識装置へ送信する通信手段と、を備えたことを特徴とする音声入力端末器が提供される。

【0007】

本発明において、前記情報とは、前記音声入力端末器に関わる固有の情報又は利用されている周辺環境や利用者自体に関わる利用状況の情報であって、例えば

、音声入力のためのマイクの特性等の前記音声入力端末器自体の機能の特性、周辺の雑音特性等の、該音声入力端末器が用いられている周辺の環境情報、若しくは、その音声入力端末器を用いる話者の話者特性等が含まれる。また、この情報には、例えば、前記入力手段から得た当該情報の元データに音響分析処理等を施したものも含まれる。

【0008】

本発明の音声入力端末器においては、前記音声認識装置から、前記情報に基づく通信のためのデータ変換条件を受け取った場合に、当該変換条件に基づいて前記音声データを変換する手段を備えることもできる。

【0009】

また、本発明の音声入力端末器においては、前記情報を保存する手段と、各通信時において前記情報に変更があったか否かを判定する手段と、前記情報に変更がなかった場合は、その旨を前記音声認識装置へ通知する手段と、を備えることもできる。

【0010】

また、本発明の音声入力端末器においては、前記情報に基づいて、音声認識モデルを作成する手段を備え、前記通信手段は、前記情報及び／又は当該音声認識モデルを、前記音声認識装置へ送信することもできる。

【0011】

また、本発明によれば、有線又は無線の通信網を介して音声入力端末器から伝送される音声データに対して音声認識処理を実行する音声認識装置であって、前記音声入力端末器から、音声認識のための前記音声入力端末器固有の又は利用状況の情報を受け取った場合に、当該情報に基づいて音声認識処理を実行する手段を備えたことを特徴とする音声認識装置が提供される。

【0012】

また、本発明によれば、有線又は無線の通信網を介して音声入力端末器から伝送される音声データに対して音声認識処理を実行する音声認識装置であって、伝送された音声データに基づいて、音声認識のための前記音声入力端末器固有の又は利用状況の情報を作成する手段と、当該情報に基づいて音声認識処理を実行す

る手段と、を備えたことを特徴とする音声認識装置が提供される。

【 0 0 1 3 】

本発明の音声認識装置においては、前記情報に基づいて音声認識モデルを作成する手段を備えることもできる。

【 0 0 1 4 】

また、本発明によれば、有線又は無線の通信網を介して音声入力端末器から伝送される音声データに対して音声認識処理を実行する音声認識装置であって、前記音声入力端末器から、音声認識のための前記音声入力端末器固有の又は利用状況の情報を受け取った場合に、前記情報に基づいて、通信のためのデータ変換条件を決定する手段と、当該データ変換条件を前記音声入力端末器へ送出する手段と、を備えたことを特徴とする音声認識装置が提供される。

【 0 0 1 5 】

また、本発明によれば、有線又は無線の通信網を介して音声入力端末器から伝送される音声データに対して音声認識処理を実行する音声認識装置であって、伝送された音声データに基づいて、音声認識のための前記音声入力端末器固有の又は利用状況の情報を作成する手段と、前記情報に基づいて、通信のためのデータ変換条件を決定する手段と、当該データ変換条件を前記音声入力端末器へ送出する手段と、を備えたことを特徴とする音声認識装置が提供される。

【 0 0 1 6 】

本発明の音声認識装置においては、前記データ変換条件が、前記情報に基づいて作成された量子化テーブルに基づくデータ変換条件を含むこともできる。

【 0 0 1 7 】

また、本発明の音声認識装置においては、前記音声入力端末器が複数の場合、各々の前記音声入力端末器毎に、前記情報を保存する手段を備えることもできる。

【 0 0 1 8 】

また、本発明の音声認識装置においては、前記音声入力端末器が複数の場合、各々の前記音声入力端末器毎に、前記音声認識モデルを保存する手段を備えることもできる。

【 0 0 1 9 】

また、本発明の音声認識装置においては、前記音声入力端末器が複数の場合、各々の前記音声入力端末器毎に、前記データ変換条件を保存する手段を備えることもできる。

【 0 0 2 0 】

また、本発明によれば、有線又は無線の通信網を介して通信可能な音声入力端末器と音声認識装置とからなる音声通信システムであって、前記音声入力端末器は、音声入力手段と、音声認識のための前記音声入力端末器固有の又は利用状況の情報を作成する手段と、当該情報を前記音声認識装置へ送信する通信手段と、を備え、前記音声認識装置は、前記情報に基づいて音声認識処理を実行する手段を備えたことを特徴とする音声通信システムが提供される。

【 0 0 2 1 】

また、本発明によれば、有線又は無線の通信網を介して通信可能な音声入力端末器と音声認識装置とからなる音声通信システムであって、前記音声認識装置は、前記音声入力端末器からの音声データに基づいて、音声認識のための前記音声入力端末器固有の又は利用状況の情報を作成する手段と、前記情報に基づいて音声認識処理を実行する手段と、を備えたことを特徴とする音声通信システムが提供される。

【 0 0 2 2 】

また、本発明によれば、有線又は無線の通信網を介して通信可能な音声入力端末器と音声認識装置とからなる音声通信システムであって、前記音声入力端末器は、音声入力手段と、音声認識のための前記音声入力端末器固有の又は利用状況の情報を作成する手段と、当該情報を前記音声認識装置へ送信する通信手段と、を備え、前記音声認識装置は、前記情報に基づいて、通信のためのデータ変換条件を決定する手段と、当該データ変換条件を前記音声入力端末器に送信する手段と、を備えたことを特徴とする音声通信システムが提供される。

【 0 0 2 3 】

また、本発明によれば、有線又は無線の通信網を介して通信可能な音声入力端末器と音声認識装置とからなる音声通信システムであって、前記音声認識装置は

、前記音声入力端末器からの音声データに基づいて、音声認識のための前記音声入力端末器固有の又は利用状況の情報を作成する手段と、前記情報に基づいて、通信のためのデータ変換条件を決定する手段と、当該データ変換条件を前記音声入力端末器に送信する手段と、を備えたことを特徴とする音声通信システムが提供される。

【 0 0 2 4 】

また、本発明によれば、有線又は無線の通信網を介して、音声入力端末器から音声認識装置へ音声データを伝送する音声通信方法であって、前記音声入力端末器において、音声認識のための前記音声入力端末器固有の又は利用状況の情報を作成する工程と、当該情報を前記音声認識装置へ送信する通信工程と、を含むことを特徴とする音声通信方法が提供される。

【 0 0 2 5 】

また、本発明によれば、有線又は無線の通信網を介して音声入力端末器から伝送される音声データに対して音声認識処理を実行する音声通信方法であって、前記音声入力端末器から、音声認識のための前記音声入力端末器固有の又は利用状況の情報を受け取った場合に、当該情報に基づいて音声認識処理を実行する工程を含むことを特徴とする音声通信方法が提供される。

【 0 0 2 6 】

また、本発明によれば、有線又は無線の通信網を介して音声入力端末器から伝送される音声データに対して音声認識処理を実行する音声通信方法であって、前記音声入力端末器から伝送されたデータに基づいて、音声認識のための前記音声入力端末器固有の又は利用状況の情報を作成する工程と、当該情報に基づいて音声認識処理を実行する工程と、を含むことを特徴とする音声通信方法が提供される。

【 0 0 2 7 】

また、本発明によれば、有線又は無線の通信網を介して音声入力端末器から伝送される音声データに対して音声認識処理を実行する音声通信方法であって、前記音声入力端末器から、音声認識のための前記音声入力端末器固有の又は利用状況の情報を受け取った場合に、当該情報に基づいて通信のためのデータ変換条件

を決定する工程と、当該データ変換条件を前記音声入力端末器へ送出する工程と、を含むことを特徴とする音声通信方法が提供される。

【 0 0 2 8 】

また、本発明によれば、有線又は無線の通信網を介して音声入力端末器から伝送される音声データに対して音声認識処理を実行する音声通信方法であって、

前記音声入力端末器から伝送されたデータに基づいて、音声認識のための前記音声入力端末器固有の又は利用状況の情報を作成する工程と、当該情報に基づいて通信のためのデータ変換条件を決定する工程と、当該データ変換条件を前記音声入力端末器へ送出する工程と、を含むことを特徴とする音声通信方法が提供される。

【 0 0 2 9 】

また、本発明によれば、有線又は無線の通信網を介して通信可能な音声入力端末器と音声認識装置との間での音声通信方法であって、前記音声入力端末器において、音声認識のための前記音声入力端末器固有の又は利用状況の情報を作成する工程と、当該情報を前記音声認識装置へ送信する通信工程と、を含み、前記音声認識装置において、前記情報に基づいて音声認識処理を実行する工程、を含むことを特徴とする音声通信方法が提供される。

【 0 0 3 0 】

また、本発明によれば、有線又は無線の通信網を介して通信可能な音声入力端末器と音声認識装置との間での音声通信方法であって、前記音声認識装置において、前記音声入力端末器からの音声データに基づいて、音声認識のための前記音声入力端末器固有の又は利用状況の情報を作成する工程と、前記情報に基づいて音声認識処理を実行する工程と、を含むことを特徴とする音声通信方法が提供される。

【 0 0 3 1 】

また、本発明によれば、有線又は無線の通信網を介して通信可能な音声入力端末器と音声認識装置との間での音声通信方法であって、前記音声入力端末器において、音声認識のための前記音声入力端末器固有の又は利用状況の情報を作成する工程と、当該情報を前記音声認識装置へ送信する通信工程と、を含み、前記音

声認識装置において、前記情報に基づいて、通信のためのデータ変換条件を決定する工程と、当該データ変換条件を前記音声入力端末器に送信する工程と、を含むことを特徴とする音声通信方法が提供される。

【0032】

また、本発明によれば、有線又は無線の通信網を介して通信可能な音声入力端末器と音声認識装置との間での音声通信方法であって、前記音声認識装置において、前記音声入力端末器からの音声データに基づいて、音声認識のための前記音声入力端末器固有の又は利用状況の情報を作成する工程と、前記情報に基づいて、通信のためのデータ変換条件を決定する工程と、当該データ変換条件を前記音声入力端末器に送信する工程と、を含むことを特徴とする音声通信方法が提供される。

【0033】

また、本発明によれば、有線又は無線の通信網を介して音声入力端末器から音声認識装置へ音声データを伝送するために、コンピュータを、音声認識のための前記音声入力端末器固有の又は利用状況の情報を作成する手段、当該情報を前記音声認識装置へ送信する通信手段、として機能させるプログラムを記録した記憶媒体が提供される。

【0034】

また、本発明によれば、有線又は無線の通信網を介して音声入力端末器からの音声データについて音声認識処理を実行するために、コンピュータを、前記音声入力端末器から、音声認識のための前記音声入力端末器固有の又は利用状況の情報を受け取った場合に、当該情報に基づいて音声認識処理を実行する手段、として機能させるプログラムを記録した記憶媒体が提供される。

【0035】

また、本発明によれば、有線又は無線の通信網を介して音声入力端末器からの音声データについて音声認識処理を実行するために、コンピュータを、前記音声入力端末器から伝送されたデータに基づいて、音声認識のための前記音声入力端末器固有の又は利用状況の情報を作成する手段、当該情報に基づいて音声認識処理を実行する手段、として機能させるプログラムを記録した記憶媒体が提供され

る。

【 0 0 3 6 】

また、本発明によれば、有線又は無線の通信網を介して音声入力端末器からの音声データについて音声認識処理を実行するために、コンピュータを、前記音声入力端末器から、音声認識のための前記音声入力端末器固有の又は利用状況の情報を受け取った場合に、当該情報に基づいて通信のためのデータ変換条件を決定する手段、当該データ変換条件を前記音声入力端末器へ送出する手段、として機能させるプログラムを記録した記憶媒体が提供される。

【 0 0 3 7 】

また、本発明によれば、有線又は無線の通信網を介して音声入力端末器からの音声データについて音声認識処理を実行するために、コンピュータを、前記音声入力端末器から伝送されたデータに基づいて、音声認識のための前記音声入力端末器固有の又は利用状況の情報を作成する手段、当該情報に基づいて通信のためのデータ変換条件を決定する手段、当該データ変換条件を前記音声入力端末器へ送出する手段、として機能させるプログラムを記録した記憶媒体が提供される。

【 0 0 3 8 】

【発明の実施の形態】

以下、本発明の好適な実施の形態について、添付図面を参照して説明する。

【 0 0 3 9 】

図 1 は、本発明の一実施形態に係る音声通信システムの構成図を示した図である。

【 0 0 4 0 】

音声通信システムは、音声入力端末器としての携帯端末 1 0 0 と、音声認識装置としての本体 2 0 0 と、これらを通信可能に接続する通信回線 3 0 0 と、からなる。

【 0 0 4 1 】

携帯端末 1 0 0 は、音声の入出力を行う入出力部 1 0 1 と、本体 2 0 0 との通信処理を実行する通信制御部 1 0 2 と、入力された音声に対して音響処理を行う音響処理部 1 0 3 と、携帯端末 1 0 0 の固有の又は利用状況の情報（以下、本実

施形態では環境情報という。)を作成する環境情報生成部104と、音声通信情報生成部105と、を備える。

【0042】

本体200は、携帯端末100の環境情報に基づく処理を行う環境適応部201と、携帯端末100との通信処理を実行する通信制御部202と、携帯端末100からの音声データに対して音声認識処理を実行する音声認識部203と、通信のためのデータ変換条件を設定する音声通信情報生成部204と、音声認識モデル保持部205と、アプリケーション206と、を備える。

【0043】

次に、係る構成からなる音声通信システムの動作手順について図2を参照して説明する。図2は、音声通信システムの処理のフローチャートである。

【0044】

音声通信システムの処理は、環境情報を解析する初期設定モードと、音声データの通信を行う音声認識モードと、からなる。

【0045】

ステップS401では、全ての処理の開始を示している。入出力部101から処理開始のための情報を通信制御部102を通じて、本体200の通信制御部202に送られる。

【0046】

ステップS402では、選択的に、音声認識部203あるいはアプリケーション206から、メッセージが携帯端末100に送られる。例えば、環境情報に基づく教師付き話者適応を行なう場合には、発声内容のリストを送り、携帯端末100の入出力部101からメッセージ（音声又は文字）として出力する。また、環境情報に基づくマイクロホン適応を行う場合には、数秒の音声を発声してもらう旨を携帯端末100の入出力部101からメッセージとして出力する場合もある。一方、環境情報に基づく雑音適応を行う場合には、このステップS402をスキップする場合もある。

【0047】

ステップS403では、携帯端末100において環境情報を生成するために、

入出力部 101 から音声データ（雑音も含む）を取り込む。

【0048】

ステップ S404 では、取り込んだ音声データについて音響処理部 103 で音響分析を行う。なお、環境情報をモデル（平均や分散、音素モデル）に変換する場合には、環境情報生成部 104 に送る。モデルにしない場合には、通信制御部 102 から本体へ音響分析の結果を送る。なお、音響分析をせずに、音声データを直接本体に送り、本体 200 側で分析等を行ってもよい。

【0049】

ステップ S404 で環境情報をモデルに変換する場合は、ステップ S405 へ進み、環境情報生成部 104 において環境情報の生成を行なう。たとえば、雑音適応を目的とする場合には、環境情報は、非音声区間を検出して、その区間の平均と分散を求めることにより生成する。また、マイクロホン適応を目的とする場合には、環境情報は、音声区間の平均と分散を求めることにより生成する。更に、話者適応を目的とする場合には、音素モデルなどを作成する。

【0050】

ステップ S406 では、作成された環境情報のモデル或いは音響分析結果若しくは音声、本体 200 へ 102 通信制御部から送出する。

【0051】

ステップ S407 では、本体 200 が送られてきた環境情報を通信制御部 202 を介して受け取る。

【0052】

ステップ S408 では、環境情報に基づいて音声認識モデル保持部 205 の音声認識モデルについて環境適応部 201 で環境適応を行ない、新たな音声認識モデルを環境適応音声認識モデルとして更新し、音声認識モデル保持部 205 にて保持する。

【0053】

環境適応の方法としては、雑音適応の場合には、例えば、雑音モデルと音声認識モデルとから環境適応音声認識モデルを作成する PMC 法などを用いることができる。マイクロホン適応の場合では、例えば、適応用音声の平均と音声認識モ

デルとを用いて、環境適応音声認識モデルを作成するCMS法などを用いることができる。

【0054】

また、話者適応の場合には、例えば、話者適応モデルと音声認識モデルを用いて話者適応モデルを作成する方法等を用いることができる。さらに、環境情報がモデルではなく音声や音響分析結果で送られてきた場合には、本体200側で環境情報をモデル化してさらに適応する方法も可能である。また、直接、音声や音響分析結果を用いて環境適応する方法、EM学習方法やVFS話者適応方法などあらゆる方法が環境適応方法として可能となる。環境適応した音声認識モデルを作成することにより、認識性能を向上させることが可能となる。

【0055】

なお、音声認識モデルを携帯端末100側で作成し、これを本体200へ送出して用いることができることも言うまでもない。

【0056】

ステップS409では、音声認識の通信効率を向上させるために、通信用の音声情報を作成するテーブルの環境適応を204音声通信情報生成部にて行なう。

ここでは、環境適応された音声認識モデルの分布を用いて、音声認識に用いるパラメータの各次元のスカラー量子化テーブルを作成する方法として説明する。

この方法にはいろいろな方法が考えられるが、もっとも簡単な方法は、各次元全体の3 σ の中から最大値と最小値を探し、その間を等分割する方法である。

【0057】

さらに、量子点数を減らすためには、全分布を一分布にマージし、その3 σ （例えば、ガウス分布において出現するサンプルのほとんどが含まれる範囲）の最大値、最小値を探し、その間を等分割する方法もある。

【0058】

或いは、さらに精密には、全分布の片寄りにあわせて、量子点を割り振る方法などが考えられる。本方法では、環境適応された音声認識モデルの分布を用いて各次元のスカラー量子化テーブルを作成するため、認識性能を落すことなく、通信のためのbit rateを落すことが可能になり、効率的な通信が出来る。

【0 0 5 9】

ステップ S 4 1 0 では、作成したスカラー量子化テーブルを携帯端末 1 0 0 に送信する。

【0 0 6 0】

ステップ S 4 1 1 では、作成されたスカラー量子化テーブルを携帯端末 1 0 0 で受け取り、音声通信情報生成部 1 0 5 に格納する。

【0 0 6 1】

以上により初期設定モードが終了する。なお、携帯端末 1 0 0 が複数存在する場合には、本体 2 0 0 は、環境情報や、音声認識モデル、若しくは、量子化テーブル等のデータを各携帯端末毎に保存することもできる。

【0 0 6 2】

次に音声認識モードに移行する。

【0 0 6 3】

ステップ S 4 1 2 では、入出力部 1 0 1 から音声を入力する。

【0 0 6 4】

ステップ S 4 1 3 では、入力された音声データを音響処理部 1 0 3 で音響分析し、音声通信情報生成部 1 0 5 に送る。

【0 0 6 5】

ステップ S 4 1 4 では、音声通信情報生成部 1 0 5 において、スカラー量子化テーブルを用いて、音声データの音響分析結果をスカラー量子化し、音声通信情報として符合化する。更に、符号化された音響データを本体 2 0 0 に通信制御部 1 0 2 を介して送信する。

【0 0 6 6】

ステップ S 4 1 5 では、本体 2 0 0 において、音声認識部 2 0 3 で、受信した音声データを複合化し、音声認識処理を実行し、その認識結果を出力する。この際、音声認識処理においては、先に作成した音声認識モデルを用いることは言うまでもない。

【0 0 6 7】

ステップ S 4 1 6 では、音声認識の結果がアプリケーション 2 0 6 で解釈され

、その結果に応じたアプリケーションを求め、通信制御部 2 0 2 へ送る。

【 0 0 6 8 】

ステップ S 4 1 7 では、アプリケーション結果を、本体 2 0 0 の通信制御部 2 0 2 を介して携帯端末 1 0 0 へ送出する。

【 0 0 6 9 】

ステップ S 4 1 8 では、携帯端末 1 0 0 が、通信制御部 1 0 2 を介してアプリケーション結果を受け取る。

【 0 0 7 0 】

ステップ S 4 1 9 では、携帯端末 1 0 0 が、入出力部 1 0 1 よりアプリケーション結果を出力する。音声認識を継続する際は、ステップ S 4 1 2 にもどる。

【 0 0 7 1 】

ステップ S 4 2 0 では、通信を終了する。

【 0 0 7 2 】

このように、本実施形態の音声通信システムでは、携帯端末 1 0 0 の環境情報に適應した音声認識モデルにより音声認識を行うので、各々の携帯端末器に対応して最適な音声認識を実行することができる。また、環境情報に基づいて通信条件を設定するので、各々の携帯端末器に対応して通信効率の向上を図ることができる。

【 0 0 7 3 】

なお、本実施形態では、雑音の場合には、雑音区間のパラメータの平均と分散を求め、それを本体に送り、音声認識モデルを P M C 法により雑音適應を行う方法を例示したが、その他の雑音適應方法も採用し得ることはいうまでもない。また、端末側で平均と分散を求め、送信する方法で記述してあるが、音声情報を送り、本体側で平均と分散を求め、雑音適應を行うこともできる。

【 0 0 7 4 】

また、本実施形態では、マイク特性の場合には、ある適当な時間の音声区間のパラメータの平均と分散を求め、それを本体に送り、音声認識モデルを C M S 法によりマイク特性適應を行う方法を例示したが、その他のマイク特性適應方法も採用し得ることはいうまでもない。また、端末側で平均と分散を求め、送信する

方法で記述してあるが、音声情報を送り、本体側で平均と分散を求め、雑音適応を行うこともできる。

【0075】

また、本実施形態では、話者適応法の場合には、事前に簡単な話者性を表す音素モデルを作成し、それを本体に送り、音声認識モデルを話者適応する方法により行う方法を例示したが、音声情報を送り、本体側で音声を用いて話者適応を行うこともできる。この場合にも、他の様々な話者適応法を採用し得ることもいうまでもない。

【0076】

また、本実施形態では、雑音適応、マイクロホン適応、話者適応に関して独立に記してあるが、これらを適宜組み合わせて採用することもできる。

【0077】

また、本実施形態では、音声認識モードに先だって初期設定モードを行うこととしているが、一旦初期設定モードが完了すれば、同じ環境にて音声認識を再開する場合には、いきなり音声認識モードから開始することもできる。この場合は、携帯端末100側において、先の環境情報を記憶しておき、音声認識を再開する際に作成した環境情報と比較して、変化がなければその旨の通知を本体200側へ送ることにより、若しくは、送出された環境情報等に基づいて本体200側で判断するようにすることにより、実行することができる。

【0078】

また、本実施形態では、環境情報を音声認識処理と通信効率の向上のための処理との双方に用いているが、当該環境情報によりいずれか一方のみを実行してもよいことは言うまでもない。

【0079】

以上、本発明の好適な実施の形態について説明したが、本発明の目的は、前述した実施形態の機能を実現するソフトウェアのプログラムコードを記録した記憶媒体（または記録媒体）を、システムあるいは装置に供給し、そのシステムあるいは装置のコンピュータ（またはCPUやMPU）が記憶媒体に格納されたプログラムコードを読み出し実行することによっても、達成されることは言うまでもない。

この場合、記憶媒体から読み出されたプログラムコード自体が前述した実施形態の機能を実現することになり、そのプログラムコードを記憶した記憶媒体は本発明を構成することになる。また、コンピュータが読み出したプログラムコードを実行することにより、前述した実施形態の機能が実現されるだけでなく、そのプログラムコードの指示に基づき、コンピュータ上で稼働しているオペレーティングシステム(OS)などが実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

【0080】

さらに、記憶媒体から読み出されたプログラムコードが、コンピュータに挿入された機能拡張カードやコンピュータに接続された機能拡張ユニットに備わるメモリに書込まれた後、そのプログラムコードの指示に基づき、その機能拡張カードや機能拡張ユニットに備わるCPUなどが実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

【0081】

【発明の効果】

以上述べた通り、本発明によれば、音声入力端末器に関する事情に基づいて、最適な音声認識又は通信を達成することができる。

【図面の簡単な説明】

【図1】

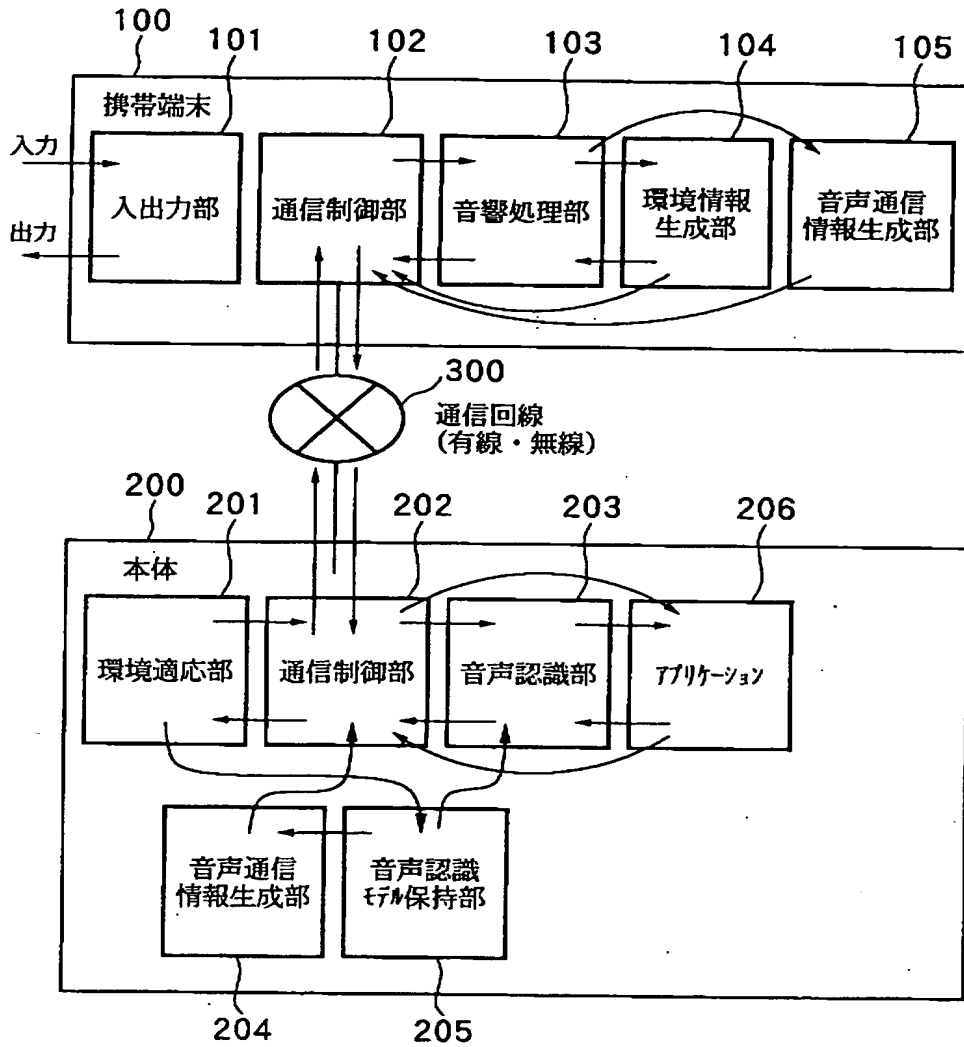
本発明の一実施形態に係る音声通信システムの構成図を示した図である。

【図2】

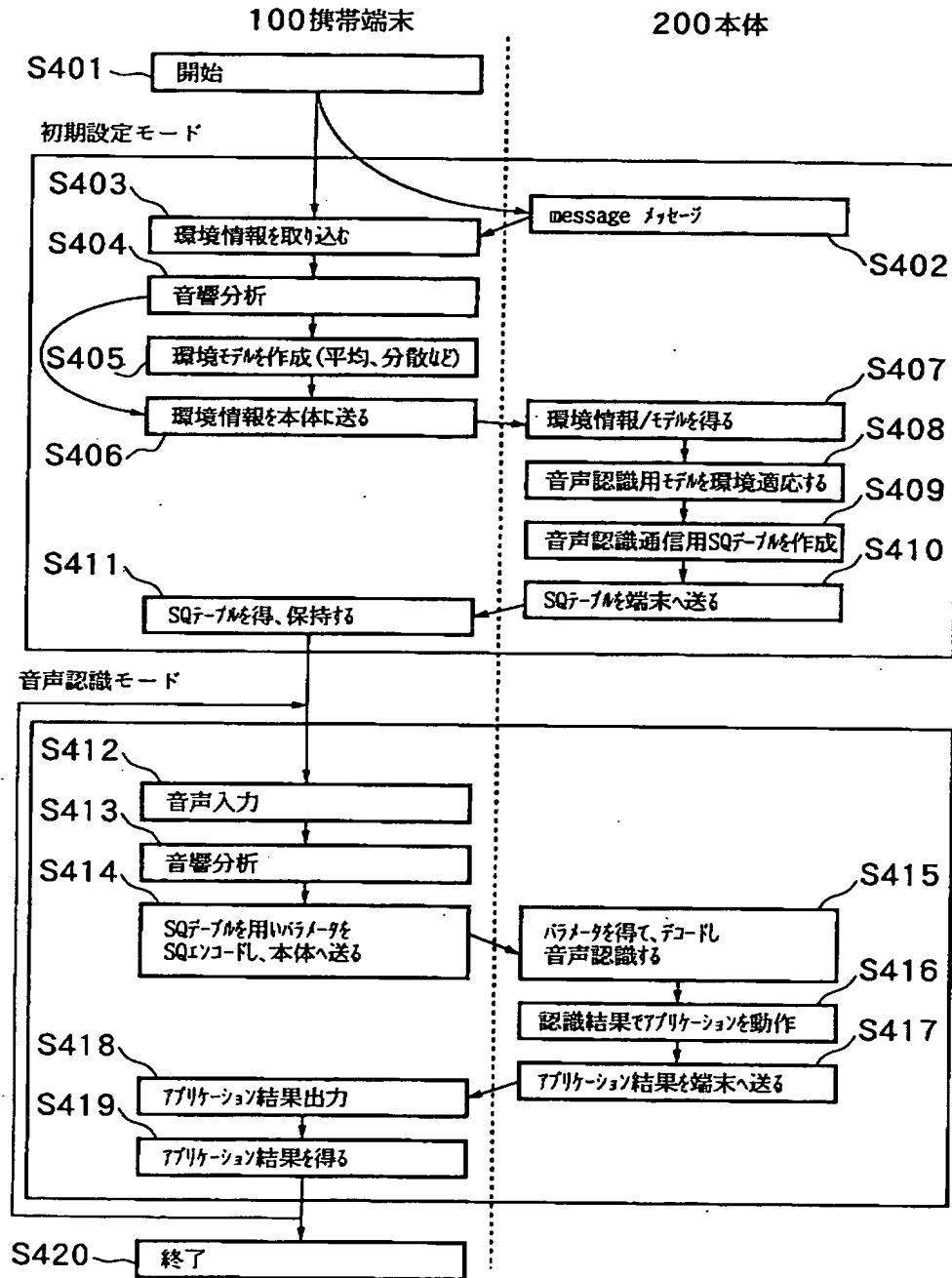
本発明の一実施形態に係る音声通信システムの処理のフローチャートである。

【書類名】 図面

【図 1】



【図 2】



【書類名】 要約書

【要約】

【課題】 音声入力端末器に関する事情に基づいて、最適な音声認識又は通信を達成し得る音声入力端末器、音声認識装置、音声通信システム、及び、音声通信方法を提供すること。

【解決手段】 初期モードとして、ステップ S 4 0 3 で、携帯端末 1 0 0 の環境情報（例えば、雑音特性、話者特性等）を取り込み、これを本体 2 0 0 へ送出する（ステップ S 4 0 6）。

本体 2 0 0 では、環境情報に基づいて音声認識又は通信時のデータ変換条件を定め（ステップ S 4 0 8 及び S 4 0 9）、以後の音声データの通信時に用いる。

【選択図】 図 2

出 願 人 履 歴 情 報

識別番号 [000001007]

1. 変更年月日 1990年 8月30日
[変更理由] 新規登録
住 所 東京都大田区下丸子3丁目30番2号
氏 名 キヤノン株式会社